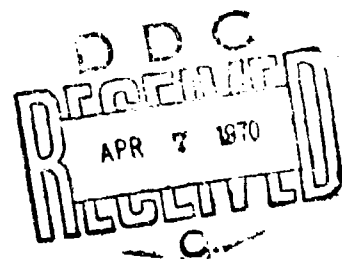
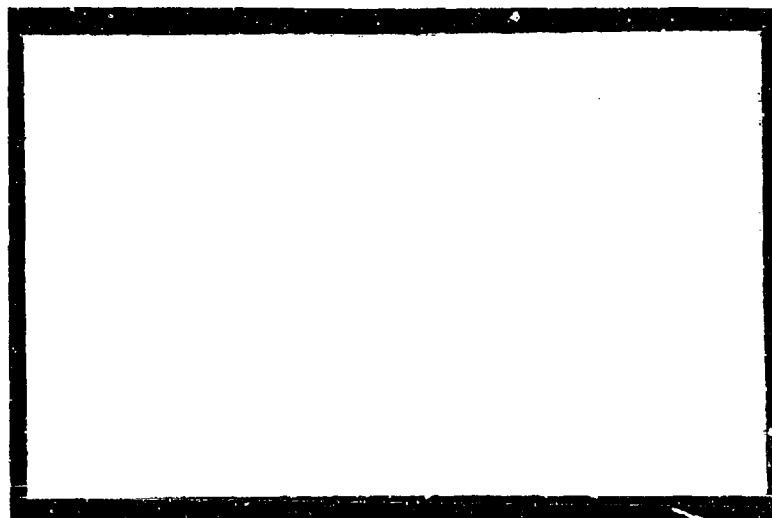
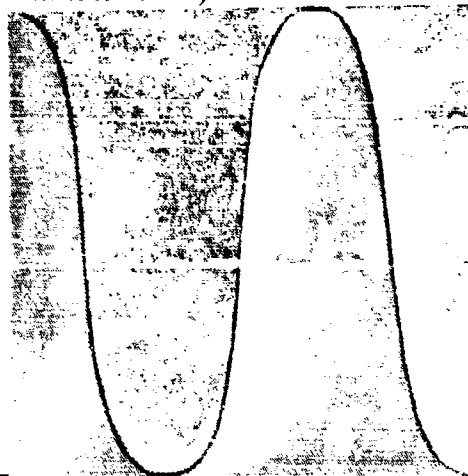


AD703191

THE UNIVERSITY  
OF WISCONSIN  
*madison, wisconsin*



UNITED STATES ARMY

MATHEMATICS RESEARCH CENTER

Reproduced by the  
CLEARINGHOUSE  
for Federal Scientific & Technical  
Information Springfield, Va. 22161



THE UNIVERSITY OF WISCONSIN  
MATHEMATICS RESEARCH CENTER, UNITED STATES ARMY

Contract No. : DA-31-124-ARO-D-462

ON THE APPROXIMATE SOLUTION OF  
FREDHOLM INTEGRAL EQUATIONS OF  
THE FIRST KIND

Grace Wahba

This document has been approved for public  
release and sale; its distribution is unlimited.

MRC Technical Summary Report # 990  
May 1969

Madison, Wisconsin 53706

# ABSTRACT

We consider methods for the approximate solution of a Fredholm integral equation of the first kind, namely

$$u(t) = \int_S K(t,s) z(s) ds, \quad t \in T$$

where  $K(t,s)$  is a given kernel,  $u(t)$  is given for  $t = t_1, t_2, \dots, t_n$ , and  $z$ , to be approximated, is an element of a reproducing kernel Hilbert space  $H_R$ .

We assume that the linear functionals  $\Lambda_t$ ,  $t \in T$  defined by

$$\Lambda_t z = \int_S K(t,s) z(s) ds$$

are all continuous on  $H_R$  and we treat the problem of approximating  $N_S z = z(s)$  for  $s \in S$  as a problem in approximating the continuous linear functional  $N_S$  by the continuous linear functionals

$$\{\Lambda_t, t = t_1, t_2, \dots, t_n\}.$$

With this point of view we do the following:

- i) give formulae for approximate solutions, under various conditions
- ii) give pointwise error bounds for these solutions
- iii) give several examples of practical reproducing kernel Hilbert spaces
- iv) show that the method of regularization of Tihonov and Glasko and the method of statistical estimation of Westwater and Strand are (equivalent) special cases of the theory presented here
- v) show how to choose the regularization parameter when errors due either to quadratures or experimental measurement is non-negligible.

ON THE APPROXIMATE SOLUTION OF FREDHOLM INTEGRAL EQUATIONS OF  
THE FIRST KIND

by

Grace Wahba

1. Introduction

1.1 Statement and Some History of the Problem

We are interested in the numerical solution of a Fredholm integral equation of the first kind, namely

$$u(t) = \int_S K(t,s) z(s) ds \quad t \in T \quad (1.1)$$

where  $S, T$  are intervals,  $K(t,s)$  is a given kernel on  $T \times S$  with appropriate properties and  $u(t)$  is given for  $t = t_1, t_2, \dots, t_n, t_i \in T$ .

It has been noted by a number of authors (see for example [7], [11], [14]) that replacing the integral in (1.1) by a quadrature approximation and inverting the resulting matrix does not always give satisfactory results. No matter how large  $n$  is, there are many functions  $z$ , including highly oscillatory ones, such that

$$u(t_i) = \int_S K(t_i,s) z(s) ds, \quad i = 1, 2, \dots, n \quad (1.2).$$

A discrete approximation of one of these functions is obtained by the matrix inversion

---

Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462 and by the Air Force Office of Scientific Research Office of Aerospace Research, United States Air Force under AFOSR Grant No. AF-AFOSR-69-1803.

technique. Which one is actually obtained, apparently depends more on the quadrature procedure than anything else. Two methods in the literature seem to have resulted in satisfactory numerical examples. The first method, called the method of regularization of Tihonov [12] [13] and studied experimentally by Tihonov and Glasko [14], goes as follows:

Let

$$M_m^\alpha(z, u) = \int_T (\hat{u}(t) - u(t))^2 dt + \alpha \int_S (L_m z(s))^2 ds \quad (1.3a)$$

where  $u$  is a given function in  $L_2(T)$ ,  $\hat{u}$ , depending on  $z$ , is given by

$$\hat{u}(t) = \int_S K(t, s) z(s) ds, \quad (1.3b)$$

$\alpha > 0$ , and  $L_m$  is an  $m^{\text{th}}$  order linear differential operator with continuous positive coefficients. Assume  $K(t, s)$  is continuous, and that

$$0 = \int_S K(t, s) z(s) ds, \quad t \in T \quad (1.4)$$

implies that  $z(s) \equiv 0$ ,  $s \in S$ .

Then, it is shown in [13], that, for every fixed  $u \in L_2[T]$ , and every  $\alpha > 0$ , there exists a unique  $2m$  times differentiable function  $\hat{z} = \hat{z}_\alpha$  which minimizes  $M_m^\alpha(z, u)$ .

Tihonov and Glasko [14] provide an argument that  $L_m$  defined by

$$L_m z = z' \quad (1.5)$$

fits into the general theory if  $K > 0$  on  $T \times S$ . They then experiment with a numerical algorithm based on finding  $\hat{z}$  to minimize  $M_m^\alpha(z, u)$  as follows. Quadrature points  $\{s_i\}_{i=1}^k$ ,  $s_i \in S$ , and  $\{t_j\}_{j=1}^n$ ,  $t_j \in T$  are chosen. Let  $t_0$  be the left boundary of  $T$ , define

$$\bar{z} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_k)'$$

$$\bar{u} = (u(t_1), u(t_2), \dots, u(t_n))'$$

and  $\bar{z}_0 = 0$ .  $\bar{u}$  is a given vector. Define  $\bar{M}_1^\alpha(\bar{z}, \bar{u})$  as

$$\begin{aligned} \bar{M}_1^\alpha(\bar{z}, \bar{u}) = & \sum_{j=1}^n d_j \left( \sum_{i=1}^k K(t_j, s_i) \sigma_i \bar{z}_i - u(t_j) \right)^2 \\ & + \alpha \sum_{i=0}^{k-1} \frac{(\bar{z}_{i+1} - \bar{z}_i)^2}{(s_{i+1} - s_i)^2} \end{aligned} \quad (1.6)$$

where  $\{d_j\}_{j=1}^n$ ,  $d_j > 0$ , and  $\{\sigma_i\}_{i=1}^k$  are appropriately chosen quadrature coefficients.  $\bar{M}_1^\alpha(\bar{z}, \bar{u})$  is to be thought of as a discretized version of  $M_1^\alpha(z, u)$  of (1.3a) with  $L_m$  given by

(1.5). Let the  $n \times k$  matrix  $\bar{K}$ , the  $n \times n$  (diagonal) matrix  $D$  and the  $k \times k$  symmetric matrix  $Q$  be defined by setting

$$\bar{M}_1^\alpha(\bar{z}, \bar{u}) = (\bar{K}\bar{z} - \bar{u})' D (\bar{K}\bar{z} - \bar{u}) + \alpha \bar{z}' Q \bar{z} \quad (1.7)$$

and matching the coefficients of  $\bar{z}$ ,  $\bar{u}$  and  $\alpha \bar{z}$  in (1.6) and (1.7). For a given  $\bar{u}$ , the vector  $\hat{\bar{z}}$  which minimizes  $\bar{M}_1^\alpha(\bar{z}, \bar{u})$  is well known to be given by

$$\hat{\bar{z}} = (\bar{K}' D \bar{K} + \alpha Q)^{-1} \bar{K}' D \bar{u} \quad (1.8)$$

It can be verified by direct manipulations upon the matrices involved that another formula for  $\hat{\bar{z}}$  is

$$\hat{\bar{z}} = Q^{-1} \bar{K}' (\bar{K} Q^{-1} \bar{K}' + \alpha D^{-1})^{-1} \bar{u} \quad (1.9)$$

Tihonov and Glasko compute  $\hat{\bar{z}}$  from (1.8) as a discrete approximation to the function  $\hat{z}_\alpha(s)$  which minimizes  $M_1^\alpha(z, u)$ . They do this experimentally, for several values of  $\alpha$ , by beginning with a  $\bar{u}$  obtained from a known function  $u(t)$  satisfying

$$u(t) = \int_S K(t, s) z^*(s) ds$$

where  $z^*$  is a given (smooth) experimental function. For certain values of  $\alpha$ , the results are "good", i.e. the components  $\hat{\bar{z}}_1$  of  $\hat{\bar{z}}$  satisfy

$$\hat{z}_1 \sim z^*(s_1).$$

More recently, Ribiere [8] has also studied the method of regularization. In [7], [8] and [14] the experimental results suggest that there is an optimum choice of  $\alpha$ . However, a theory making precise the optimal choice is apparently not available.

The second method, discussed by Strand and Westwater [11] and called "statistical estimation" of the solution, is as follows: Let the  $n \times k$  matrix  $\bar{K}$  be defined as before. Let  $Z = (z_1, z_2, \dots, z_k)'$  and  $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)'$  be normally distributed zero mean random vectors with (prior) covariance

$$EZZ' = Q^{-1}$$

$$E\epsilon\epsilon' = \alpha D^{-1}$$

$$EZ\epsilon' = 0$$

Let  $U = (U_1, U_2, \dots, U_n)$  be a normally distributed random vector defined by

$$U = \bar{K}Z + \epsilon \quad (1.10)$$

It is well known that



$$E(Z|U=\bar{u}) = Q^{-1}\bar{K}'(\bar{K}Q^{-1}\bar{K}'+\alpha D^{-1})^{-1}\bar{u} \quad (1.11)$$

Strand and Westwater replace the equation (1.1) with the model (1.10), where the unknown vector  $\bar{z}$  is replaced by the random vector  $Z$  with prior distribution  $N(0, Q^{-1})$ , and  $\epsilon$  is to be thought of as a "noise" vector independent of  $z$  and having the distribution  $N(0, \alpha D^{-1})$ . The estimate for the "random" vector  $Z$  given the "data"  $\bar{u}$  is taken as the right hand side of (1.9). Hence the method of regularization and that of statistical estimation are in practice the same for appropriate choice of  $D$  and  $Q$ . Strand and Westwater performed numerical experiments (again beginning with a smooth known solution  $z^*$ ) analogous to those of Tihonov and Glasko, and again obtained "good" results.

## 1.2 Purpose of This Note and Basic Assumptions

Both of these methods can be embedded in the general theory of the approximation of continuous linear functionals in a reproducing kernel Hilbert space. The overall purpose of this note is to demonstrate this statement in some considerable practical detail. As a byproduct of this theory we obtain a rationale for choosing  $\alpha$ , to minimize a certain error bound. We also obtain pointwise error bounds on the solution involving a Hilbert space norm of the (unknown) solution  $z$ . These results provide a criteria for the selection

of a good Hilbert space in which to operate. Several practical examples are given.

We assume the function  $z$  to be an element of a reproducing kernel Hilbert space  $\mathcal{H}$  of real-valued functions defined on  $S$ . This means that the linear functionals  $N_s$ ,  $s \in S$ , defined by

$$N_s z = z(s) \quad , \quad z \in \mathcal{H} \quad (1.12)$$

are continuous for every  $s \in S$ . This is just what is needed to establish point wise error bounds. In this case, for  $\forall s \in S$  there exists an element  $\delta_s \in \mathcal{H}$  for which

$$\langle \delta_s, z \rangle = z(s) \quad , \quad \forall z \in \mathcal{H} \quad (1.13)$$

If we define the kernel  $R(s, s')$  on  $S \times S$  as

$$\langle \delta_s, \delta_{s'} \rangle = R(s, s') \quad (1.14)$$

then  $R(s, s')$  is positive definite. Let  $R_{s_0}$  be that function defined on  $S$  whose value at  $s$  is given by

$$R_{s_0}(s) = R(s_0, s) \quad (1.15)$$

By the Moore-Aronszajn Theorem (see [2]), to every positive definite kernel  $R(s, s')$  defined on  $S \times S$  there corresponds a unique Hilbert space  $\mathcal{H} = \mathcal{H}_R$  with the following properties:

$$1) R_{s_0} \in M \quad \forall s_0 \in S$$

$$2) \langle z, R_{s_0} \rangle = z(s_0) \quad , \quad z \in M \quad , \quad s_0 \in S$$

Hence  $\delta_s$  of (1.13) is given by

$$\delta_s = R_s \quad (1.16)$$

and

$$\langle R_s, R_{s'} \rangle = R(s, s') \quad (1.17)$$

Equation (1.17) is the source of the terminology "reproducing kernel". The elements  $\{R_s, s \in S\}$  clearly span  $M$ .

We make sufficient assumptions on  $K$  to ensure that the linear functionals  $\Lambda_t$  defined on  $M_R$  by

$$\Lambda_t z = \int_S K(t, s) z(s) ds \quad , \quad z \in M_R \quad (1.18)$$

are continuous for every fixed  $t \in T$ . The problem of estimating  $z(s)$  for a particular  $s \in S$  given that  $z \in M_R$  and the information of (1.1), namely,  $\Lambda_{t_i} z = u(t_i)$ ,  $i = 1, 2, \dots, n$  may then be viewed as that of approximating the continuous linear functional  $N_s$  defined by (1.12) by the continuous linear functionals  $\{\Lambda_{t_i}\}_{i=1}^n$  defined by (1.18).

To discuss the problem practically, we must know when  $\Lambda_t$  defined by (1.18) is continuous, and how to find its representer  $\eta_t$ . The answer is given by Theorem (1.1) and its corollary which will conclude Section 1.

Once the representers  $\{\eta_{t_i}\}_{i=1}^n$  are known, there are several points of view that one may adopt when approximating one continuous linear functional by several others. Perhaps the simplest is as follows. Let  $\#_n$  be the subspace spanned by  $\{\eta_{t_i}\}_{i=1}^n$ , and  $P_n$  the projection operator onto this subspace. Suppose  $\#_n$  is  $n$  dimensional. Approximate  $R_s$  by  $P_n R_s$ . Then if

$$\langle \eta_{t_i}, z \rangle = u(t_i), \quad i = 1, 2, \dots, n$$

we have

$$\begin{aligned} \langle P_n R_s, z \rangle &= \langle R_s, P_n z \rangle = (P_n z)(s) \\ &= (\eta_{t_1}(s), \eta_{t_2}(s), \dots, \eta_{t_n}(s)) \begin{pmatrix} \langle \eta_{t_1}, \eta_{t_1} \rangle & \dots & \langle \eta_{t_1}, \eta_{t_n} \rangle \\ \vdots & & \vdots \\ \langle \eta_{t_n}, \eta_{t_1} \rangle & \dots & \langle \eta_{t_n}, \eta_{t_n} \rangle \end{pmatrix}^{-1} \begin{pmatrix} u(t_1) \\ \vdots \\ u(t_n) \end{pmatrix} \end{aligned}$$

$P_n z$  is, of course, that element in  $\#$  of minimum norm which satisfies (1.2). One then has the point wise error bound

$$|z(s) - (P_n z)(s)| = |\langle z - P_n z, R_s \rangle| = |\langle z - P_n z, R_s - P_n R_s \rangle|$$

$$\leq \|z - P_n z\| \|R_s - P_n R_s\| \leq \|z\| \|R_s - P_n R_s\| \quad (1.19a)$$

where  $\|R_s - P_n R_s\|$  can be computed for any fixed  $s$ . Hence in solving any particular practical problem, one should expect good results if one can choose a reproducing kernel Hilbert space in which the norm of the unknown function  $z$  is small. We actually generalize this point of view in two ways. First, if  $\mathcal{H}_1$  is a Hilbert subspace of  $\mathcal{H}$  of codimension  $m < n$ , then it is possible to approximate  $R_s$  by  $\{n_{t_i}\}_{i=1}^n$  in such a way that  $R_s$  and its approximant differ by an element in  $\mathcal{H}_1$ . In this case, it is shown (in Lemma 2.5) that the right hand side of (1.19a) may be replaced by

$$\|P_1 z\| \|P_1(R_s - \hat{R}_s)\| \quad (1.19b)$$

where  $P_1$  is the projection operator onto  $\mathcal{H}_1$ , and  $\hat{R}_s$  is the approximating element. Here we will be able to calculate  $\|P_1(R_s - \hat{R}_s)\|$ , and one chooses an  $\mathcal{H}$  and an  $\mathcal{H}_1$  such that it is believed that the norm in  $\mathcal{H}$  of  $P_1 z$  is small. This is the geometry of the approximation of continuous linear functionals used by Golomb and Weinberger [4]. For other somewhat related work on the approximation of one continuous linear functional by several others see [3] and [9].

Further generalizing, we actually consider in some detail the cases where  $u(t_i)$  are known only up to some error (experimental error) or the values  $\eta_{t_i}(s)$  of the representers  $\eta_{t_i}$  are known only imperfectly (quadrature error). In this case the error bounds analogous to (1.19a) involve the magnitude of the (experimental or quadrature) error. We show how an approximate minimization of these error bounds leads in a natural way to a family of algorithms, which contains the method(s) of regularization and statistical estimation.

Section 2 provides statements of all the Hilbert space lemmas that are used in the sequel. They are generally well known. Since they may be proved by elementary methods, no proofs are provided. Section 3 provides examples of some reproducing kernel Hilbert spaces. Concrete examples are fairly hard to come by. Section 4 presents the main theorems concerning approximations to the solution  $z$  and their properties. For concreteness sake, but without loss of generality, the theorems are stated with respect to a particular reproducing kernel Hilbert space which is related to the method of regularization as discussed in [12], [13], [14]. Section 5 discusses the introduction of quadrature error and appropriate ways of dealing with it, and a rationale for choosing  $\alpha$  is given. In Section 6, the precise relationship between

statistical estimation, and the approximation of continuous linear functionals is described, and it is shown how in general one obtains the same numerical result from the two approaches.

We conclude this Section with Theorem 1.1 and its corollary.

### 1.3 A Preliminary Theorem

Theorem 1.1. Let  $\Lambda$  be a continuous linear functional on  $\mathcal{H}_R$ , and let  $\psi \in \mathcal{H}_R$  be defined by

$$\Lambda z = \langle \psi, z \rangle \quad , \quad z \in \mathcal{H}_R$$

Then  $\psi(s)$ ,  $s \in S$ , is given by the formula:

$$\psi(s) = \Lambda R_s \quad (1.20)$$

Conversely, let  $\tilde{\Lambda}$  be a linear functional defined on  $\mathcal{H}_R$  and suppose that the function  $\tilde{\psi}$  defined in  $S$  by

$$\tilde{\psi}(s) = \tilde{\Lambda} R_s \quad (1.21)$$

satisfies  $\tilde{\psi} \in \mathcal{H}_R$ . Then  $\tilde{\Lambda}$  is continuous.

Proof. Since  $\psi(s) \in \mathcal{H}_R$ ,  $\psi(s) = \langle \psi, R_s \rangle = \Lambda R_s$ . Conversely, if  $\tilde{\psi} \in \mathcal{H}_R$ , then the continuous linear functional  $\tilde{\Lambda}$  defined by

$$\tilde{\Lambda}z = \langle \tilde{\psi}, z \rangle \quad (1.22)$$

coincides with  $\tilde{\Lambda}$  on the span of  $\{R_s, s \in S\}$ , since  $\tilde{\Lambda}R_s = \tilde{\psi}(s)$ . But the span of  $\{R_s, s \in S\}$  is  $\mathcal{H}_R$ .

Corollary. Let  $K$  be such that

$$\int_S \int_S K(t, s) R(s, s') K(t, s') ds ds' = \theta^2(t) \quad (1.23)$$

is well defined and finite as a Riemann integral for each fixed  $t \in T$ . Then the linear functional  $\Lambda_t$  defined by

$$\Lambda_t z = \int_S K(t, u) z(u) du, \quad z \in \mathcal{H}_R \quad (1.24)$$

is continuous on  $\mathcal{H}_R$  and

$$\Lambda_t z = \langle \eta_t, z \rangle, \quad t \in T, \quad z \in \mathcal{H}_R$$

where  $\eta_t \in \mathcal{H}_R$  is defined by

$$\eta_t(s) = \Lambda_t R_s = \int_S K(t, u) R_s(u) du, \quad (1.25)$$

and  $||\Lambda_t||^2 = \theta^2(t)$ .

Proof. The hypotheses on  $K$  imply that there exists (for each fixed  $t$ ), a triangular array



$$\begin{array}{c}
s_{11} \\
s_{12}, s_{22} \\
\vdots \\
s_{1k}, s_{2k}, \dots, s_{kk} \\
\vdots
\end{array}$$

such that the Riemann sums

$$\sum_{i=1}^{k-1} \sum_{j=1}^{l-1} K(t, s_{ik}) R(s_{ik}, s_{jl}) K(t, s_{jl}) (s_{i+1,k} - s_{ik}) (s_{j+1,l} - s_{jl})$$

converge to  $\theta^2(t)$ , as  $k, l \rightarrow \infty$ ,  $\max |s_{i+1,k} - s_{ik}| \rightarrow 0$ . But then, using (1.17), it follows that the sequence  $\eta_t^{(k)}$ ,  $k = 1, 2, \dots$ , defined by

$$\eta_t^{(k)} = \sum_{i=1}^{k-1} K(t, s_{ik}) (s_{i+1,k} - s_{ik}) R_{s_{ik}} \quad k = 1, 2, \dots$$

is a Cauchy sequence in  $\mathcal{H}_R$ ,  $\|\eta_t^{(k)}\|^2 \rightarrow \theta^2(t)$ , and  $\eta_t^{(k)}$  converges pointwise (in  $s$ ) to  $\eta_t$  given by (1.25), which is therefore in  $\mathcal{H}_R$ .

## 2. Hilbert Space Lemmas

In the lemmas of this section we have the following:  
 $m \leq n$ ,  $\mathcal{H}$  is a Hilbert space,

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$$

where  $\mathcal{H}_0$  is an  $m$ -dimensional subspace,  $\mathcal{H}_1 = \mathcal{H}_0^\perp$  and  $P_0$  and  $P_1$  are the projection operators onto  $\mathcal{H}_0$  and  $\mathcal{H}_1$  respectively.  $\{\eta_i\}_{i=1}^n$  are, in each lemma,  $n$  elements in  $\mathcal{H}$  satisfying the following two conditions

- (i)  $\{P_0 \eta_i\}_{i=1}^n$  span  $\mathcal{H}_0$
- (ii)  $\{P_1 \eta_i\}_{i=1}^n$  are linearly independent.

Let

$$P_1 \eta_i = \xi_i, \quad i = 1, 2, \dots, n$$

let  $\{\phi_\nu\}_{\nu=1}^m$  be a (fixed) orthonormal basis for  $\mathcal{H}_0$ , and let  $X$  and  $[ ]$  be the  $m \times n$  and  $n \times n$  matrices with entries defined by

$$\begin{aligned} [X]_{\mu i} &= \langle \phi_\mu, \eta_i \rangle, \quad \mu = 1, 2, \dots, m \\ [ ]_{ij} &= \langle \xi_i, \xi_j \rangle, \quad i, j = 1, 2, \dots, n \end{aligned} \tag{2.1}$$

Conditions (i) and (ii) guarantee that  $X$  and  $[ ]$  are of full rank. We let  $\phi$ ,  $\xi$  and  $\eta$  be vectors of  $m$ ,  $n$  and  $n$  elements of  $\mathcal{H}$  respectively, given by

$$\phi = (\phi_1, \phi_2, \dots, \phi_m)$$

$$\xi = (\xi_1, \xi_2, \dots, \xi_n)$$

$$\eta = (\eta_1, \eta_2, \dots, \eta_n)$$

and  $\bar{u}$  be a given vector of real numbers,

$$\bar{u} = (u_1, u_2, \dots, u_n).$$

Lemma 2.1. The solution to the problem: Find  $z \in \#$  to minimize  $\langle P_1 z, P_1 z \rangle$  subject to  $\langle z, \eta_i \rangle = u_i$ ,  $i = 1, 2, \dots, n$ , is unique, and is given by  $\hat{z}$ ,

$$\hat{z} = \phi(X \bar{\Gamma}^{-1} X')^{-1} X \bar{\Gamma}^{-1} \bar{u} + \xi (\bar{\Gamma}^{-1} - \bar{\Gamma}^{-1} X' (X \bar{\Gamma}^{-1} X')^{-1} X \bar{\Gamma}^{-1}) \bar{r}. \quad (2.2)$$

Lemma 2.2. Let  $\delta_0 \in \#$  be given. The solution to the problem: Find  $y \in \#$  of the form

$$y = \sum_{i=1}^n c_i \eta_i \quad (2.3)$$

where  $\bar{c} = (c_1, c_2, \dots, c_n)$  is a vector of real numbers, to minimize

$$||P_1(\delta_0 - y)||^2 \quad (2.4)$$

subject to

$$||P_0(\delta_0 - y)||^2 = 0 \quad (2.5)$$

is unique and is given by  $\hat{\delta}_0$ ,

$$\hat{\delta}_0 = \bar{\phi} (X \bar{\Sigma}^{-1} X')^{-1} X \bar{\Sigma}^{-1} \eta' + \bar{\xi} (\bar{\Sigma}^{-1} - \bar{\Sigma}^{-1} X' (X \bar{\Sigma}^{-1} X')^{-1} X \bar{\Sigma}^{-1}) \eta' \quad (2.6)$$

where  $\bar{\phi}$  and  $\bar{\xi}$  are the vectors of real numbers (depending on  $\delta_0$ ), given by

$$\bar{\phi} = (\langle \delta_0, \phi_1 \rangle, \langle \delta_0, \phi_2 \rangle, \dots, \langle \delta_0, \phi_m \rangle) \quad (2.7)$$

$$\bar{\xi} = (\langle \delta_0, \xi_1 \rangle, \langle \delta_0, \xi_2 \rangle, \dots, \dots, \langle \delta_0, \xi_n \rangle).$$

Let the bounded linear operators  $A$  and  $A^*$  be defined by (2.2) and (2.6) respectively as

$$Az = \hat{z}, \quad \text{if } u_i = \langle z, \eta_i \rangle, \quad i = 1, 2, \dots, n, z \in \mathcal{H}$$

$$A^* \delta_0 = \hat{\delta}_0, \quad \delta_0 \in \mathcal{H}$$

Inspection of (2.2) and (2.6) reveals that  $A$  and  $A^*$  are each idempotent, and are adjoint to each other, that is:

Lemma 2.3 If  $z$  is of the form of the right hand side of (2.2) for some  $\bar{u}$ , then  $z = \hat{z}$ .

Lemma 2.4

$$\langle \hat{z}, \delta_0 \rangle = \langle z, \hat{\delta}_0 \rangle = \langle \hat{z}, \hat{\delta}_0 \rangle, \quad z, \delta_0 \in \mathcal{H} \quad (2.8)$$

Lemma 2.5

$$\begin{aligned} \langle z - \hat{z}, \delta_0 \rangle^2 &= \langle z - \hat{z}, \delta_0 - \hat{\delta}_0 \rangle^2 \leq \|P_1(z - \hat{z})\| \|P_1(\delta_0 - \hat{\delta}_0)\|^2 \\ &\leq \|P_1 z\|^2 \|P_1(\delta_0 - \hat{\delta}_0)\|^2 \quad (2.9) \end{aligned}$$

Furthermore, a calculation gives

Lemma 2.6

$$\begin{aligned} \|P_1(\delta_0 - \hat{\delta}_0)\|^2 &= \langle P_1 \delta_0, P_1 \delta_0 \rangle - \bar{\xi} \bar{\xi}^{-1} \bar{\xi} \\ &\quad + (\bar{\phi} - \bar{\xi} \bar{\xi}^{-1} X') (X \bar{\xi}^{-1} X')^{-1} (\bar{\phi} - \bar{\xi} \bar{\xi}^{-1} X'), \end{aligned} \quad (2.10)$$

where  $\bar{\phi}$  and  $\bar{\xi}$  are given by (2.7).

Lemma 2.7 Let  $V = \{v_{ij}\}$  be a non-negative definite  $n \times n$  matrix with  $ij^{\text{th}}$  entry of  $V^{-1}$  given by  $v^{ij}$ . Then the solution to the problem: Find  $z \in \mathbb{R}^n$  to minimize

$$M(z) = \sum_{i,j=1}^n (\langle z, \eta_i \rangle - u_i) v^{ij} (\langle z, \eta_j \rangle - u_j) + \langle P_1 z, P_1 z \rangle \quad (2.11)$$

is unique, and is given by  $\tilde{z}$ ,

$$\tilde{z} = \phi (XS^{-1}X')^{-1} XS^{-1} \bar{u} + \xi (S^{-1} - S^{-1}X' (XS^{-1}X')^{-1} XS^{-1}) \bar{u}. \quad (2.12)$$

where

$$S = \bar{J} + V. \quad (2.13)$$

Lemma 2.8 Let  $\{\epsilon_i\}_{i=1}^n$  be  $n$  elements in  $\mathcal{H}$  satisfying  $\langle \epsilon_i, \epsilon_j \rangle = v_{ij}$ , and let  $\delta_0 \in \mathcal{H}$  be given. The solution to the problem: Find  $y \in \mathcal{H}$  of the form

$$y = \sum_{i=1}^n d_i (\eta_i + \epsilon_i) \quad (2.14)$$

where  $\bar{d} = (d_1, d_2, \dots, d_n)$  is a vector of real numbers to be found, to minimize

$$\|\delta_0 - \sum_{i=1}^n d_i \eta_i\|^2 + \|\sum_{i=1}^n d_i \epsilon_i\|^2 \quad (2.15)$$

subject to

$$\|P_0(\delta_0 - \sum_{i=1}^n d_i \eta_i)\|^2 = 0 \quad (2.16)$$

is unique, and is given by  $\tilde{\delta}_0$ ,

$$\tilde{\delta}_0 = \bar{\phi}(XS^{-1}X')^{-1}XS^{-1}(\eta + \epsilon)' + \bar{\xi}(S^{-1} - S^{-1}S'(XS^{-1}X')^{-1}XS^{-1})(\eta + \epsilon)', \quad (2.17)$$

where  $\bar{\phi}$ ,  $\bar{\xi}$ , depending on  $\delta_0$  are given by (2.7),  $S$  is given by (2.13), and the vector of functions  $(\eta + \epsilon)$  is given by

$$\eta + \epsilon = (\eta_1 + \epsilon_1, \eta_2 + \epsilon_2, \dots, \eta_n + \epsilon_n) \quad (2.18)$$

Lemma 2.9 Let  $z$  satisfy  $\langle z, \eta_1 + \epsilon_1 \rangle = u_1, i = 1, 2, \dots, n$ , then if  $\tilde{z}$  is given by (2.12) and  $\tilde{\delta}_0$  by (2.17), then

$$\langle z, \tilde{\delta}_0 \rangle = \langle \tilde{z}, \delta_0 \rangle \quad (2.19)$$

Lemma 2.10 For  $\bar{d} = (d_1, d_2, \dots, d_n)$  any vector of real numbers, and  $z$  and  $\tilde{z}$  as in Lemma 2.9,  $\tilde{\delta}_0$  as in (2.17)

$$\langle z - \tilde{z}, \delta_0 \rangle^2 = \langle z, \delta_0 - \tilde{\delta}_0 \rangle^2 \leq 2[||z||^2 (||\delta_0 - \sum_{i=1}^n d_i \eta_i||^2 + ||\sum_{i=1}^n d_i \epsilon_i||^2)] \quad (2.20)$$

If  $\epsilon_i \in \mathcal{H}_1, i = 1, 2, \dots, n$ , then  $\delta_0 - \tilde{\delta}_0 \in \mathcal{H}_1$ , and

$$\langle z - \tilde{z}, \delta_0 \rangle^2 \leq 2[||P_1 z||^2 (||\delta_0 - \sum_{i=1}^n d_i \eta_i||^2 + ||\sum_{i=1}^n d_i \epsilon_i||^2)] \quad (2.21)$$

### 3. Examples of Reproducing Kernel Hilbert Spaces

#### 3.1 Generalities

Let  $S, \Omega$  each be the real line or a closed subset, let  $G(s, u)$  be defined on  $S \times \Omega$  with the property that  $G(s, u) \in L_2(\Omega)$  for every fixed  $s$ , and suppose further that

$$0 = \int_{\Omega} G(s,u)p(u)du, \quad p \in L_2(\Omega), \quad s \in S \quad (3.1)$$

implies that  $p = 0$ . Then the range of the operator  $G$  defined by

$$(Gp)(s) = \int_{\Omega} G(s,u)p(u)du \quad (3.2)$$

is a reproducing kernel Hilbert space of functions defined on  $S$ . If we call this space  $\mathcal{H}_1$ , then it has reproducing kernel  $R_1$  given by

$$R_1(s,s') = \int_{\Omega} G(s,u)G(s',u)du \quad (3.3)$$

with inner product

$$\langle f_1, f_2 \rangle = \int_{\Omega} p_1(u)p_2(u)du \quad (3.4)$$

where

$$f_i(s) = (Gp_i)(s) = \int_{\Omega} G(s,u)p_i(u)du, \quad i = 1, 2 \quad (3.5)$$

Suppose  $\mathcal{H}_1$  is contained in some larger Hilbert space of functions defined on  $S$ , and  $\{\phi_v\}_{v=1}^m$  are  $m$  orthonormal functions on  $S$  all in  $\mathcal{H}_1^\perp$ . Then  $\mathcal{H}_0$ , the  $m$  dimensional space spanned by  $\{\phi_v\}_{v=1}^m$ , is a reproducing kernel space with



reproducing kernel  $R_0$  given by

$$R_0(s, s') = \sum_{v=1}^m \phi_v(s) \phi_v(s') \quad (3.6)$$

and  $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$  is a reproducing kernel space with reproducing kernel  $R$  given by

$$R(s, s') = R_0(s, s') + R_1(s, s') \quad (3.7)$$

### 3.2 Reproducing Kernel Spaces Associated with Green's Functions

Let  $L_m$  be an  $m^{\text{th}}$  order linear differential operator on  $[0, 1]$  with  $m$ -dimensional null space. Let the linear functionals  $M_\mu$  be defined by

$$M_\mu f = \sum_{v=1}^m \gamma_{\mu v} f^{(v-1)}(0), \quad \mu = 1, 2, \dots, m \quad (3.8)$$

where the matrix  $\{\gamma_{\mu v}\}$  is non-singular. Let  $G(s, u)$  be the Green's function for the equation  $L_m f = g$ ,  $M_\mu f = 0$ ,  $\mu = 1, 2, \dots, m$ , and let  $\{\phi_v\}_{v=1}^m$  span the null space of  $L_m$  with

$$M_\mu \phi_v = \delta_{\mu v} \quad (3.9)$$

where  $\delta_{\mu v}$  is the Kronecker  $\delta$ . We may take  $\mathcal{H}_1$  as the collection  $\{f: L_m f \in L_2[0, 1], M_\mu f = 0, \mu = 1, 2, \dots, m\}$  and  $\mathcal{H}_0$

the space spanned by  $\{\phi_v\}_{v=1}^m$ .  $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$  then has inner product

$$\langle f_1, f_2 \rangle = \sum_{v=1}^m (M_v f_1) (M_v f_2) + \int_0^1 (L_m f_1)_s (L_m f_2)_s ds \quad (3.10a)$$

and

$$\|P_1 f\|^2 = \int_0^1 (L_m f)_s^2 ds. \quad (3.10b)$$

Two simple examples are:

a)  $L_m = \frac{d^m}{ds^m}$

$$G(s, u) = \frac{(s-u)^{m-1}}{(m-1)!} \quad (3.11)$$

$$\phi_v(s) = \frac{s^{v-1}}{(v-1)!}; \quad v = 1, 2, \dots, m$$

$$\gamma_{\mu v} = 1, \mu = v, = 0 \text{ otherwise}$$

b)  $L_m = \prod_{v=1}^m (D + \alpha_v)$ ,  $\{\alpha_v\}_{v=1}^m$  distinct positive real numbers

$$G(s, u) = \sum_{v=1}^m c_v e^{-\alpha_v(s-u)}, \quad s \geq u, \quad c_v^{-1} = \prod_{j \neq v} (\alpha_j - \alpha_v)$$

$$= 0 \quad \text{otherwise} \quad (3.12)$$

$$\phi_v(s) = c_v e^{-\alpha_v s}$$

$$M_v f = \prod_{l \neq v} (D + \alpha_l)$$

where  $(D+\alpha_\nu)f = f' + \alpha_\nu f$ . This type of reproducing kernel space may be used to construct spline functions. (See [3])

### 3.3 Spaces of Functions with Rapidly Decreasing Fourier Transform

Let  $W(\lambda)$  be a complex-valued Hermitian function of  $\lambda$  on  $(-\infty, \infty)$  with  $W(\lambda) \in L_2(-\infty, \infty)$  and  $|W(\lambda)| > 0$ , all real  $\lambda$ . Let  $G(\tau)$  be the inverse Fourier-Plancherel Transform of  $W(\lambda)$  which we write as

$$G(\tau) = \int_{-\infty}^{\infty} e^{-i\tau\lambda} W(\lambda) d\lambda \quad (3.13)$$

Let

$$R(s, s') = \int_{-\infty}^{\infty} G(s-u)G(s'-u)du = \int_{-\infty}^{\infty} e^{-i(s-t)\lambda} W(\lambda)W^*(\lambda) d\lambda \quad (3.14)$$

$R(s, s')$  is positive definite, and, since  $|W(\lambda)| > 0$ ,

$$0 = \int_{-\infty}^{\infty} G(s-u)p(u)du, \quad -\infty < s < \infty, p \in L_2 \quad (3.15)$$

implies  $p(u) = 0$ . # , the reproducing kernel space for  $R(s, s')$  is the collection of all functions  $f$  of the form

$$f(s) = \int_{-\infty}^{\infty} G(s-u)p(u)du \quad p \in L_2(-\infty, \infty) \quad (3.16)$$

with inner product (by Parseval's theorem), given by

$$\langle f_1, f_2 \rangle = \int_{-\infty}^{\infty} \frac{F_1(\lambda) F_2^*(\lambda)}{W(\lambda) W^*(\lambda)} d\lambda \quad (3.17)$$

where  $F_i$  is the Fourier-Plancherel Transform of  $f_i$ .  $H$  contains only functions whose Fourier transforms  $F(\lambda)$  decrease sufficiently rapidly that  $F(\lambda)/W(\lambda) \in L_2(-\infty, \infty)$ .

For a concrete example, let  $q, p$  be non-negative integers with  $q < p$ , let  $\{\alpha_\mu\}_{\mu=1}^p, \{\beta_\nu\}_{\nu=1}^q$  be  $p + q$  distinct positive real numbers, let

$$\begin{aligned} P(\lambda) &= \prod_{\mu=1}^p (i\lambda - \alpha_\mu) \\ Q(\lambda) &= \prod_{\nu=1}^q (i\lambda - \beta_\nu) \end{aligned} \quad (3.18)$$

$$W(\lambda) = Q(\lambda)/P(\lambda).$$

We have

$$\begin{aligned} G(\tau) &= \sum_{\mu=1}^p c_\mu e^{-\alpha_\mu \tau}, \quad \tau \geq 0 \\ &= 0 \end{aligned} \quad (3.19)$$

where

$$c_\mu = \prod_{k=1}^q (\beta_k - \alpha_\mu) / \prod_{j \neq \mu} (\alpha_j - \alpha_\mu)$$

and

$$R(s, s') = \sum_{v=1}^p \theta_v e^{-\alpha_v |s-s'|} \quad (3.20)$$

$$\theta_v = \sum_{\mu=1}^p \frac{c_\mu c_v}{(\alpha_\mu + \alpha_v)}$$

(for  $q = 0$ ,  $p = 1$  we have  $R(s, s') = e^{-\alpha |s-s'|}$ ).

Letting  $m = p - q$ ,  $\mathcal{M}$  is here the collection of all functions with absolutely continuous  $m$ -th derivative and square integrable  $m^{\text{th}}$  derivative. Let

$$G_Q(\tau) = \sum_{v=1}^q d_v e^{-\beta_v \tau}, \quad \tau \geq 0$$

$$= 0 \quad (3.21)$$

where

$$d_v^{-1} = \prod_{j \neq v} (\beta_j - \beta_v).$$

Let

$$(G_Q f)_t = \int_{-\infty}^{\infty} G_Q(t-s) f(s) ds \quad (3.22)$$

and let

$$L_p f = \prod_{\mu=1}^p (D + \alpha_\mu) f. \quad (3.23)$$

For  $f \in \mathcal{H}$ ,  $(G_Q f)$  has  $p-1$  absolutely continuous and  $p$  square integrable derivatives, and the inner product in  $\mathcal{H}$  is given by

$$\langle f_1, f_2 \rangle = \int_{-\infty}^{\infty} (L_p G_Q f_1)_t (L_p G_Q f_2)_t dt. \quad (3.24)$$

### 3.4 Spaces of Band Limited Functions

Let  $\mathcal{F}_\Lambda$  be the subset of  $L_2(-\infty, \infty)$  of functions whose Fourier transforms vanish outside  $[-\Lambda, \Lambda]$ .

Let

$$G(\tau) = \int_{-\infty}^{\infty} s(\tau-u) w(u) du \quad (3.25)$$

where

$$s(\tau) = \frac{\sin \Lambda \tau}{\pi \tau} \quad (3.26)$$

$$w(u) = \int_{-\Lambda}^{\Lambda} e^{iu\lambda} W(\lambda) d\lambda$$

and  $W(\lambda)$  is assumed to be a Hermitian function of  $\lambda$  with no real zeroes in  $[-\Lambda, \Lambda]$ . The equation

$$0 = \int_{-\infty}^{\infty} G(t-u) p(u) du \quad -\infty < t < \infty \quad (3.27)$$

has solutions in  $L_2(-\infty, \infty)$  but not in  $\mathfrak{H}_\Lambda$ , since (3.25)-(3.27) imply  $W(\lambda)P(\lambda) = 0$ ,  $-\Lambda \leq \lambda \leq \Lambda$ , where  $P(\lambda)$  is the Fourier-Plancherel transform of  $p(u)$ . (Note that  $S(\lambda)$ , the Fourier-Plancherel transform of  $\sin \Lambda \tau / \pi \tau$  is  $S(\lambda) = 1$ ,  $|\lambda| \leq \Lambda$ ,  $= 0$  otherwise.)

$\mathfrak{H}_\Lambda$  is a reproducing kernel Hilbert space with inner product

$$\langle f_1, f_2 \rangle = \int_{-\Lambda}^{\Lambda} \frac{F_1(\lambda) F_2(\lambda)}{W(\lambda) W^*(\lambda)} d\lambda \quad (3.28)$$

where  $F_i$  is the Fourier-Plancherel transform of  $f_i$ ,  $i = 1, 2$ .

The reproducing kernel is

$$R(s, s') = \int_{-\Lambda}^{\Lambda} e^{-i\lambda(s-s')} W(\lambda) W^*(\lambda) d\lambda. \quad (3.29)$$

If we set

$$W(\lambda)^2 = 1, \quad |\lambda| \leq \Lambda \quad (3.30)$$

$$= 0 \quad \text{otherwise}$$

then

$$R(s, s') = \frac{\sin \Lambda(s-s')}{\pi(s-s')} \quad (3.31)$$

and  $\mathfrak{H}_\Lambda$  is a Hilbert subspace of  $L_2(-\infty, \infty)$ . Slepian and Pollack [10] consider the space of bandlimited functions with reproducing kernel (3.31). A restatement of the sampling

theorem tells us that the functions  $\eta_j$  defined by

$$\eta_j(s) = \frac{1}{\sqrt{2\Lambda}} \int_{-\Lambda}^{\Lambda} e^{i(s-2\pi j/\Lambda)} w(\lambda) d\lambda, \quad j = \dots -1, 0, 1, \dots$$

(3.32)

are a complete orthonormal basis for  $\mathfrak{F}_\Lambda$

#### 4. Typical Theorems

A variety of theorems now fall out by applying Theorem 1.1 and the lemmas of Section 2 to the reproducing kernel Hilbert spaces of Section 3. For concreteness only we state them with reference to the example of Section 3.2, since this example provides a direct comparison with the method of regularization as discussed by Tihonov and Glasko.

Thus, let  $L_m$  be an  $m^{\text{th}}$  order linear differential operator on  $[0,1]$  with  $m$ -dimensional null space spanned by  $\{\phi_v\}_{v=1}^m$ , and  $\mathcal{A} = \{z: L_m z \in L_2[0,1], z^{(m-1)} \text{ absolutely continuous}\}$ . Let  $G(s,u)$  be the Green's function for the equation

$$L_m f = g, \quad f^{(v)}(0) = 0, \quad v = 0, 1, 2, \dots, m-1 \quad \underline{1}]$$

---

1] The reader may verify that the choice of the matrix  $\{\lambda_{\mu\nu}\}$  in (3.8) is irrelevant to the solution of our problems, we take it as  $I$ .



and

$$R_1(s, s') = \int_0^1 G(s, u) G(s', u) du \quad (4.1)$$

$$R(s, s') = \sum_{v=1}^m \phi_v(s) \phi_v(s') + \int_0^1 G(s, u) G(s', u) du \quad (4.2)$$

We suppose that  $K(t, s)$  satisfies that

$$\int_0^1 \int_0^1 K(t_i, s) R(s, s') K(t_i, s') ds ds' \quad i = 1, 2, \dots, n$$

is well defined and finite as a Riemann integral.

Let  $\eta_i (\eta_i \in \mathcal{M})$  be defined by

$$\eta_i(s) = \int_0^1 K(t_i, u) R(s, u) du \quad (4.3)$$

Define the elements  $\xi_i \in \mathcal{M}$ , and the matrices  $X$  and  $\Gamma$  by

$$P_i \eta_i = \xi_i, \quad \xi_i(s) = \int_0^1 K(t_i, u) R_i(s, u) du, \quad i = 1, 2, \dots, n, \quad (4.4)$$

$$X = \{\chi_{\mu i}\}, \mu = 1, 2, \dots, n, i = 1, 2, \dots, n$$

$$\Gamma = \{\sigma_{ij}\}, i, j = 1, 2, \dots, n$$

$$\chi_{\mu i} = \langle \eta_i, \phi_\mu \rangle = \int_0^1 K(t_i, u) \phi_\mu(u) du \quad (4.5)$$

$$\sigma_{ij} = \langle \xi_i, \xi_j \rangle = \int_0^1 \int_0^1 K(t_i, s) R_1(s, s') K(t_j, s') ds ds' \quad (4.6)$$

Theorem 4.1 Let  $\bar{u} = (u_1, u_2, \dots, u_n)$  be a vector of  $n$  real numbers, let  $V = \{v_{ij}\}$  be an  $n \times n$  symmetric positive definite matrix with  $V^{-1} = \{v^{ij}\}$ , and suppose that matrices  $X$  and  $\Gamma$  defined by (4.5) and (4.6) be of full rank. For  $z \in \mathcal{M}$ , define the numbers  $\hat{u}_i$  by

$$\hat{u}_i = \hat{u}_i(z) = \langle \eta_i, z \rangle = \int_0^1 K(t_i, s) z(s) ds \quad (4.7)$$

$$i = 1, 2, \dots, n$$

Then

(i) There is a unique solution  $\hat{z}_\alpha$  to the problem:

Find  $z \in \mathcal{M}$  to minimize

$$M(z) = \sum_{i,j=1}^n (\hat{u}_i - u_i) v^{ij} (\hat{u}_j - u_j) + \alpha \int_0^1 (L_m z)^2 ds \quad (4.8)$$

given by

$$\hat{z}_\alpha = \phi (X\Gamma_\alpha^{-1}X')^{-1}X\Gamma_\alpha^{-1}\bar{u} + \xi (\Gamma_\alpha^{-1} - \Gamma_\alpha^{-1}X'(X\Gamma_\alpha^{-1}X')^{-1}X\Gamma_\alpha^{-1})\bar{u},$$

(4.9)

where  $\phi$  and  $\xi$  are the vectors of functions

$$\phi = (\phi_1, \phi_2, \dots, \phi_m)$$

$$\xi = (\xi_1, \xi_2, \dots, \xi_n)$$

and

$$\Gamma_\alpha = \Gamma + \alpha V \quad (4.10)$$

$$(ii) \quad \lim_{\alpha \rightarrow 0} \hat{z}_\alpha = \hat{z}_0, \text{ given by}$$

$$\hat{z}_0 = \phi (X\Gamma^{-1}X')^{-1}X\Gamma^{-1}\bar{u} + \xi (\Gamma^{-1} - \Gamma^{-1}X'(X\Gamma^{-1}X')^{-1}X\Gamma^{-1})\bar{u},$$

(4.11)

is the (unique) solution to the problem: Find  $z \in H$  to minimize

$$\int_0^1 (L_m z)_s^2 ds$$

subject to  $\hat{u}_i(z) = u_i, i = 1, 2, \dots, n.$

(iii) For any  $z \in \mathbb{H}$  satisfying

$$u_i = \int_0^1 K(t_i, s) z(s) ds \quad i = 1, 2, \dots, n \quad (4.12)$$

and  $\hat{z}_0$  given by (4.11) we have the pointwise error bounds

$$|\hat{z}_0(s) - z(s)|^2 \leq \sigma^2(s) \left\{ \int_0^1 (L_m z)^2 ds \right\} \quad (4.13)$$

where

$$\sigma^2(s) = R_1(s, s) - \xi(s) \int_0^1 \xi(s)' \quad (4.14a)$$

$$+ (\bar{\phi}(s) - \bar{\xi}(s) \int_0^1 X') (X \int_0^1 X')^{-1} (\bar{\phi}(s) - \bar{\xi}(s) \int_0^1 X')$$

and

$$\bar{\phi}(s) = (\phi_1(s), \phi_2(s), \dots, \phi_m(s),$$

$$\bar{\xi}(s) = (\xi_1(s), \xi_2(s), \dots, \xi_n(s)) \quad (4.14b)$$

(iv) If  $z \in \mathbb{H}$  has the form

$$z(s) = \sum_{v=1}^m c_v \phi_v(s) + \sum_{i=1}^n d_i \xi_i(s) \quad (4.15)$$

where  $\bar{d} = (d_1, d_2, \dots, d_n)$  is any vector satisfying  $X\bar{d}' = 0$ ,

and we set  $u_i$  in (ii) as

$$u_i = \int_0^1 K(t_i, s) z(s) ds \quad (4.16)$$

then the solution  $\hat{z}_0$  in (ii) satisfies

$$\hat{z}_0 = z \quad (4.17)$$

Proof: Assertion (i) is Lemma 2.7 and (ii) is Lemma 2.1. The remarkable error bound of (iii) is obtained from Lemmas 2.5 and 2.6 upon setting  $\delta_0$  as the representer of the continuous linear functional  $N_s$  defined by  $N_s z = z(s)$ ,  $z \in H$ , that is,  $\delta_0 = R_s$ . In this case we have

$$|\hat{z}_0(s) - z(s)| = |\langle \hat{z}_0 - z, R_s \rangle|$$

$$\langle R_s, \phi_\mu \rangle = \phi_\mu(s) \quad (4.18)$$

$$\langle R_s, \xi_i \rangle = \xi_i(s)$$

(iv) is another way of writing Lemma 2.3.

Other minimization problems may be handled within the context of the geometry of reproducing kernel Hilbert space. For example, the solution to the problem: Find  $z \in H$  to minimize  $M(z)$  of (4.8) subject to the linear inequalities

$$a_l \leq z(s_l) \leq b_l, \quad l = 1, 2, \dots, k, \quad s_l \in S \quad (4.19)$$

may be reduced to a standard problem of minimizing a (finite) quadratic form subject to linear inequalities, as follows:

Inequalities (4.19) may be written

$$a_l \leq \langle R_{s_l}, z \rangle \leq b_l, \quad l = 1, 2, \dots, k \quad (4.20)$$

Then, any solution  $z$  to this minimization problem must be of the form

$$z = \sum_{v=1}^m c_v \phi_v + \sum_{i=1}^n d_i \xi_i + \sum_{l=1}^k e_l R_{s_l} \quad (4.21)$$

for some coefficients  $\{c_v\}$ ,  $\{d_i\}$  and  $\{e_l\}$ . By substituting (4.21) into (4.20) and (4.8), the minimization problem reduces to the standard problem of minimizing a quadratic form in the coefficients, subject to a set of linear inequalities. This minimization problem may be recognized as a problem in control theory. See, for example [5].

Side conditions which are just sufficient to specify  $P_0 z$  give especially simple looking answers and error bounds. For example:

Theorem 4.2. Under the assumptions of Theorem 4.1:

(i) The solution to the problem: find  $z \in \mathcal{H}$  to minimize (4.8) subject to the boundary conditions  $\theta$

$$\theta: M_v z = \langle \phi_v, z \rangle = \theta_v, \quad v = 1, 2, \dots, m \quad (4.22)$$

is unique and is given by  $\hat{z}_\alpha$ ,

$$\hat{z}_\alpha = \phi \bar{\theta}' + \xi \Gamma_\alpha^{-1} (\bar{u} - X' \bar{\theta}') \quad (4.23)$$

where  $\bar{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$

ii)  $\lim_{\alpha \rightarrow 0} \hat{z}_\alpha = \hat{z}_0$ , where

$$\hat{z}_0 = \phi \bar{\theta}' + \xi \Gamma^{-1} (\bar{u} - X' \bar{\theta}') \quad (4.24)$$

is the solution to the problem: Find  $z \in \mathcal{H}$  to minimize

$$\int_0^1 (L_m z)_s^2 ds \quad (4.25)$$

subject to

$$u_i = \int_0^1 K(t_i, s) z(s) ds, \quad i = 1, 2, \dots, n \quad (4.26)$$

$$M_v z = \theta_v, \quad v = 1, 2, \dots, m. \quad (4.22)$$

(iii) If  $z \in \mathcal{H}$  satisfies

$$u_i = \int_0^1 K(t_i, s) z(s) ds \quad (4.27a)$$

and

$$M_v z = \theta_v, \quad v = 1, 2, \dots, m \quad (4.27b)$$

then  $\hat{z}_0$  of (4.24) satisfies

$$|z(s) - \hat{z}_0(s)|^2 \leq \sigma^2(s) \left\{ \int_0^1 (L_m z)_u^2 du \right\} \quad (4.28a)$$

where

$$\sigma^2(s) = R_1(s, s) - \bar{\xi}(s) \int^{-1} \bar{\xi}(s) \quad (4.28b)$$

and  $\bar{\xi}(s)$  is given by (4.14b).

Proof:  $z \in \mathcal{H}$ ,  $M_v z = \theta_v$ ,  $v = 1, 2, \dots, m$  imply that

$$P_0 z = \sum_{v=1}^m \theta_v \phi_v$$

and

$$\int_0^1 K(t_i, s) z(s) ds = \langle \eta_i, z \rangle = \langle \xi_i, P_1 z \rangle + \sum_{v=1}^m \theta_v \chi_{vi} \quad (4.29)$$

In this case (4.8) may be written

$$M(z) = \sum_{i,j} (\langle \xi_i, P_1 z \rangle - (u_i - \sum_{v=1}^m \theta_v \chi_{vi})) v^{ij} (\langle \xi_j, P_1 z \rangle - (u_j - \sum_{v=1}^m \theta_v \chi_{vj})) \\ + \alpha \|P_1 z\| \quad (4.30)$$

Since  $\{\xi_i\}_{i=1}^n$  and  $P_1 z$  are in  $\mathcal{H}_1$ , we may find  $(P_1 z)$  to minimize (4.30) or  $(P_1 z)$  to minimize (4.25) subject to



$$u_i - \sum_{v=1}^m \theta_v x_{vi} = \langle \xi_i, P_1 z \rangle, \quad i = 1, 2, \dots, n$$

via Lemmas 2.7 and 2.1 respectively, by setting  $m = 0$ ,

$M = M_1$  in these lemmas. To prove (iii) let  $P_\xi$  be the projection operator onto the subspace spanned by  $\{\xi_i\}_{i=1}^n$ .

If  $z \in M$  satisfies (4.27a) and (4.27b), then

$$P_\xi z = P_1 \hat{z}_0 \quad (4.31a)$$

$$P_0 z = P_0 \hat{z}_0 \quad (4.31b)$$

and hence

$$z - \hat{z}_0 = (P_1 - P_\xi) z. \quad (4.32)$$

But

$$\begin{aligned} z(s) - \hat{z}_0(s) &= \langle z - \hat{z}_0, R_s \rangle \\ &= \langle (P_1 - P_\xi) z, R_s \rangle \\ &= \langle (P_1 - P_\xi) z, (P_1 - P_\xi) R_s \rangle \end{aligned} \quad (4.33)$$

Hence

$$|z(s) - \hat{z}_0(s)| \leq \| (P_1 - P_\xi) z \| \| (P_1 - P_\xi) R_s \| \leq \| P_1 z \| \| (P_1 - P_\xi) R_s \| \quad (4.34)$$

But

$$\sigma^2(s) = ||(P_1 - P_\xi)R_s||^2 \quad (4.35)$$

A direct comparison with the solution of the methods of regularization and statistical estimation obtains, as follows:

Set  $V^{-1} = D$  in (4.8) where  $D$  is the diagonal matrix of (1.7), and set  $\theta = 0$  in (4.23). Choose  $\{s_i\}_{i=1}^k$ ,  $s_i \in [0,1]$ , as in Section 1 and let

$$\bar{u} = (u_1, u_2, \dots, u_n)'$$

be a given vector of real numbers and

$$\bar{z} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_k)'$$

where

$$\bar{z}_i = \hat{z}_\alpha(s_i), \quad \hat{z}_\alpha = \xi \Gamma_\alpha^{-1} \bar{u}$$

Then

$$\bar{z} = \Xi (\bar{\Gamma} + \alpha D^{-1})^{-1} \bar{u} \quad (4.36)$$

where  $\Xi$  is the  $k \times n$  matrix with  $i, j^{\text{th}}$  entry

$$\xi_j(s_i) = \int_0^1 K(t_j, u) R_1(s_i, u) du \quad (4.37)$$

and  $\{$  is the  $n \times n$  matrix with  $i, j^{\text{th}}$  entry

$$\langle \xi_i, \xi_j \rangle = \int_0^1 \int_0^1 K(t_i, u) R_1(u, u') K(t_j, u') du du' \quad (4.38)$$

If  $\{\sigma_\ell\}_{\ell=1}^k$  are a set of suitably chosen quadrature coefficients, then we have, for purposes of comparison, approximately

$$\xi_j(s_i) \approx \sum_{\ell=1}^k K(t_j, u_\ell) R_1(s_i, u_\ell) \sigma_\ell \quad (4.39)$$

$$\langle \xi_i, \xi_j \rangle \approx \sum_{\ell=1}^k \sum_{\ell'=1}^k K(t_i, u_\ell) R_1(u_\ell, u_{\ell'}) K(t_j, u_{\ell'}) \sigma_\ell \sigma_{\ell'}, \quad (4.40)$$

Let  $\bar{K}$  be the  $n \times k$  matrix (of (1.7)), with  $i, j^{\text{th}}$  entry  $K(t_j, s_i) \sigma_i$ , and  $\bar{R}$  be the  $k \times k$  matrix with  $i, j^{\text{th}}$  entry  $R_1(s_i, s_j)$ . Then we may write (4.36) as

$$\bar{z} \approx \bar{R} \bar{K} (\bar{R} \bar{K} \bar{K}' + \alpha D^{-1})^{-1} \bar{u} \quad (4.41)$$

We may identify  $\bar{R}$  with  $Q^{-1}$  of section 1 as follows:

Suppose  $f \in \mathcal{H}_1$ , then if  $L_m f = g$ , we have

$$f(s) = \int_0^1 G(s, u) g(u) du \quad (4.42)$$

and

$$\int_0^1 (L_m f)^2 ds = \int_0^1 g^2(s) ds \quad (4.43)$$

Let  $\bar{f} = (f(s_1), f(s_2), \dots, f(s_k))'$ ,  $\bar{g} = (g(s_1), g(s_2), \dots, g(s_k))'$ ,  $\bar{G}$  be the  $k \times k$  matrix with  $ij^{\text{th}}$  entry  $G(s_i, s_j)\sigma_j$  and  $D_\sigma$  be the  $k \times k$  diagonal matrix with  $i, i^{\text{th}}$  entry  $\sigma_i$ .

Since

$$R(s, s') = \int_0^1 G(s, u)G(s', u)du, \quad (4.44)$$

we have

$$\bar{R} \approx \bar{G} D_\sigma^{-1} \bar{G}' \quad (4.45)$$

also

$$\bar{f} \approx \bar{G}g \quad (4.46)$$

$$f'Qf \approx \int_0^1 (L_m f)^2 ds = \int_0^1 g^2(s) ds \approx g'D_\sigma g \approx f'G^{-1'}D_\sigma G^{-1}f = f'\bar{R}^{-1}f \quad (4.47)$$

Some questions of convergence may be answered as follows:  
For simplicity, we consider that the boundary values  $\beta_0$  of the solution  $z$  to (1.1) are known. We may then consider, without loss of generality that  $z$  and  $\hat{z}_0$  are in  $\#_1$ .

Let the family of functions  $\xi_t \in H_1$ ,  $0 \leq t \leq 1$  be defined by

$$\xi_t(s) = \int_0^1 K(t,u) R_1(s,u) du \quad 0 \leq t \leq 1 \quad (4.48)$$

If the family of functions  $\{\xi_t, 0 \leq t \leq 1\}$  span the (separable) space  $H_1$ , then  $\sigma^2(s)$  of (4.28), will tend to 0 for each  $s$  as the set  $\{t_i\}_{i=1}^n$  becomes dense in  $[0,1]$ , by (4.35). A necessary and sufficient condition that  $\{\xi_t, 0 \leq t \leq 1\}$  span  $H_1$ , is that

$$\langle \xi_t, z \rangle = 0 \quad 0 \leq t \leq 1, \quad z \in H_1, \quad (4.49)$$

implies that  $z = 0$ . But (4.49) may be rewritten

$$\int_0^1 K(t,s) z(s) ds = 0, \quad 0 \leq t \leq 1, \quad z \in H_1 \quad (4.50)$$

implies  $z = 0$ . There does not seem to be a straightforward general way of establishing a rate at which

$$\sigma^2(s) = ||(P_1 - P_{\xi}) R_s||^2 \quad (4.35)$$

tends to zero, as, say  $\sup |t_{i+1} - t_i| \rightarrow 0$ , if indeed such a rate exists. However, results have been obtained regarding the convergence of  $|| (P_1 - P_{\xi}) \delta ||$  when  $\delta \in H_1$  is "very smooth".

Thus error rates for the pointwise approximation of very smooth  $z$ , or for the approximation of continuous linear functionals with very smooth representers, are available. These results will appear separately.

## 5. The Introduction of Quadrature Formulae and the Choice of a

If the integral

$$\xi_i(s) = \int_0^1 K(t_i, u) R_1(s, u) du \quad (4.4)$$

can be evaluated analytically at values of  $s$  for which it is desired to estimate  $z(s)$ , and  $\chi_{ui}$  and  $\sigma_{ij}$  of (4.5) and (4.6) are known exactly, and computational and experimental errors are negligible, then it is natural to estimate  $z(s)$  by  $\hat{z}_0(s)$  of Theorem 4.1 or Theorem 4.2. The purpose of this section is to study the situation where  $\{\chi_{ui}\}$ ,  $\{\sigma_{ij}\}$  and  $\{\xi_i(s)\}_{i=1}^n$  must be evaluated by quadratures, where quadrature error is the primary source of error. Let  $\{s_k\}_{k=1}^N$ ,  $s_k \in [0, 1]$ , and  $\{\omega_k\}_{k=1}^N$  be suitably chosen quadrature points and quadrature coefficients, respectively. We show that this situation leads in a natural way to estimating  $z$  by the solution to the problem: Find  $z \in H$  to minimize

$$\bar{M}(z) = \sum_{i=1}^n \left( u_i - \sum_{k=1}^N K(t_i, s_k) \omega_k z(s_k) \right)^2 + \lambda \int_0^1 (L_m z)^2 ds \quad (5.1)$$

where  $\lambda$  is chosen to approximate the mean square quadrature error. Define  $\hat{\eta}_i \in H$  by

$$\hat{\eta}_i = \sum_{k=1}^N K(t_i, s_k) \omega_k R_{s_k} \quad i = 1, 2, \dots, n \quad (5.2)$$

Letting  $R_{1,s_k}$  be that element of  $H_1$  whose value at  $s$  is given by  $R_{1,s_k}(s) = R_1(s, s_k)$ ,  $k = 1, 2, \dots, N$ , define  $\hat{\xi}_i \in H_1$  by

$$\hat{\xi}_i = \sum_{k=1}^N K(t_i, s_k) \omega_k R_{1,s_k}, \quad i = 1, 2, \dots, n \quad (5.3)$$

Then  $P_1 \hat{\eta}_i = \hat{\xi}_i$ ,  $i = 1, 2, \dots, n$ .  $\hat{\chi}_{\mu i}$  given by

$$\hat{\chi}_{\mu i} = \langle \hat{\eta}_i, \phi_\mu \rangle = \sum_{k=1}^N K(t_i, s_k) \omega_k \phi_\mu(s_k) \quad (5.4)$$

is a quadrature approximation to  $\chi_{\mu i}$ ,  $\hat{\sigma}_{ij}$  given by

$$\hat{\sigma}_{ij} = \langle \hat{\xi}_i, \hat{\xi}_j \rangle = \sum_{k=1}^N \sum_{\ell=1}^N K(t_i, s_k) \omega_k R_1(s_k, s_\ell) K(t_j, s_\ell) \omega_\ell \quad (5.5)$$

is a quadrature approximation to  $\sigma_{ij}$  and  $\hat{\xi}_i(s)$  given by

$$\hat{\xi}_i(s) = \sum_{k=1}^N K(t_i, s_k) \omega_k R_1(s, s_k) \quad (5.6)$$

is a quadrature approximation to  $\xi_i(s)$  of (4.4), for each  $s$ .

Define  $\epsilon_i \in \mathbb{H}$  by

$$\epsilon_i = \eta_i - \hat{\eta}_i \quad (5.7)$$

The problem may now be viewed as that of approximating  $z(s) = \langle R_s, z \rangle$  from the information  $u_i = \langle \hat{\eta}_i + \epsilon_i, z \rangle$ ,  $i = 1, 2, \dots, n$ , or, alternatively approximating  $R_s$  by  $\{\hat{\eta}_i + \epsilon_i\}_{i=1}^n$ ; where the  $\{\epsilon_i\}_{i=1}^n$  are unknown.

Let

$$y = \sum_{i=1}^n d_i (\hat{\eta}_i + \epsilon_i)$$

where  $\bar{d} = (d_1, d_2, \dots, d_n)$  are to be found so that  $y$  is a good approximation to  $R_s$ . If we try to choose  $\bar{d}$  to minimize  $\|R_s - y\|$  in the error bound

$$|\langle z, R_s - y \rangle| \leq \|z\| \|R_s - y\| \quad (5.8)$$

it is necessary to know  $\langle R_s, \epsilon_i \rangle = \epsilon_i(s)$ ,  $i = 1, 2, \dots, n$ , which is assumed unknown.

We will choose  $\bar{d}$  subject to the constraint

$$\|P_0(R_s - \sum_{i=1}^n d_i \hat{\eta}_i)\|^2 = 0$$



Then

$$||R_s - y||^2 \leq 2\{||P_1(R_s - \sum_{i=1}^n d_i \hat{\eta}_i)||^2 + ||\sum_{i=1}^n d_i \epsilon_i||^2\} \quad (5.9)$$

Let  $V$  be the matrix with  $i, j$ <sup>th</sup> entry  $v_{ij} = \langle \epsilon_i, \epsilon_j \rangle$  given by

$$\begin{aligned} \langle \epsilon_i, \epsilon_j \rangle &= \int_0^1 K(t_i, u) du \left\{ \int_0^1 K(t_j, v) R(u, v) dv - \sum_{k=1}^N K(t_j, s_k) R(u, s_k) \omega_k \right\} \\ &\quad - \sum_{l=1}^N K(t_i, s_l) \left\{ \int_0^1 K(t_j, v) R(s_l, v) dv - \sum_{k=1}^N K(t_j, s_k) R(s_l, s_k) \omega_k \right\} \end{aligned} \quad (5.10)$$

and let  $\hat{X}$  and  $\hat{\Sigma}$  be the  $n \times m$  and  $n \times n$  matrices with entries  $\{\hat{x}_{\mu i}\}$  and  $\{\hat{\sigma}_{ij}\}$  respectively, given by (5.4) and (5.5).

By Lemma 2.8 with  $\delta_0 = R_s$ , the solution  $\tilde{R}_s$  to the problem: Find  $y$  of the form

$$y = \sum_{i=1}^n d_i (\hat{\eta}_i + \epsilon_i) \quad (5.11)$$

to minimize

$$||P_1(R_s - \sum_{i=1}^n d_i \hat{\eta}_i)||^2 + ||\sum_{i=1}^n d_i \epsilon_i||^2 \quad (5.12)$$

subject to

$$||P_O(R_S - \sum_{i=1}^n d_i \hat{\eta}_i)||^2 = 0 \quad (5.13)$$

is given by

$$\begin{aligned} \tilde{R}_S &= \bar{\phi}(s) (\hat{X}\hat{S}^{-1}\hat{X}')^{-1} \hat{X}\hat{S}^{-1} (\hat{\eta} + \epsilon) \\ &+ \bar{\xi}(s) (\hat{S}^{-1} - \hat{S}^{-1}\hat{X}(\hat{X}\hat{S}^{-1}\hat{X}')^{-1}\hat{S}^{-1}) (\hat{\eta} + \epsilon) \end{aligned} \quad (5.14a)$$

where

$$\bar{\phi}(s) = (\phi_1(s), \phi_2(s), \dots, \phi_m(s)) \quad (5.14b)$$

$$\bar{\xi}(s) = (\hat{\xi}_1(s), \hat{\xi}_2(s), \dots, \hat{\xi}_n(s))$$

$$S = \hat{S} + V$$

and

$$\hat{\eta} + \epsilon = (\hat{\eta}_1 + \epsilon_1, \hat{\eta}_2 + \epsilon_2, \dots, \hat{\eta}_n + \epsilon_n)'$$

Any "optimal" approximation to  $R_S$  will depend on the unknown  $V$ . Thus it is desirable here to approximate  $V$  using whatever information is available. A plausible approximation is

$$V \approx \lambda I \quad (5.15)$$

where  $\lambda$  is a "guestimate" of the mean diagonal element of  $V$ ,

$$\lambda \approx \frac{1}{n} \sum_{i=1}^n \langle \varepsilon_i, \varepsilon_i \rangle, \quad (5.16)$$

this "estimate" being based on (5.10) with  $i = j$  and the properties of the quadrature formula being used.

Let  $\tilde{R}_{s,\lambda} \in \mathcal{M}$  be given by (5.14) with  $V$  replaced by  $\lambda I$ , that is,

$$\begin{aligned} \tilde{R}_{s,\lambda} = & \bar{\phi}(s) (\hat{X} \hat{S}_\lambda^{-1} \hat{X}')^{-1} \hat{X} \hat{S}_\lambda^{-1} (\hat{\eta} + \varepsilon) \\ & + \hat{\xi}(s) (\hat{S}_\lambda^{-1} - \hat{S}_\lambda^{-1} \hat{X} (\hat{X} \hat{S}_\lambda^{-1} \hat{X}')^{-1} \hat{S}_\lambda^{-1}) (\hat{\eta} + \varepsilon) \end{aligned} \quad (5.17)$$

where

$$\hat{S}_\lambda = \hat{S} + \lambda I.$$

Then an estimate  $\hat{z}_\lambda$  for  $z(s)$  is defined by

$$\hat{z}_\lambda(s) = \langle z, \tilde{R}_{s,\lambda} \rangle \quad (5.18)$$

with

$$\|z(s) - \hat{z}_\lambda(s)\| \leq \|z\| \|\tilde{R}_s - \tilde{R}_{s,\lambda}\|$$

The function  $\hat{z}_\lambda$  defined by (5.18) is in  $\mathcal{M}$  and may be written

$$\begin{aligned} \hat{z}_\lambda = & \bar{\phi}(\hat{X} \hat{S}_\lambda^{-1} \hat{X}') \hat{X} \hat{S}_\lambda^{-1} \bar{u} \\ & + \hat{\xi}(\hat{S}_\lambda^{-1} - \hat{S}_\lambda^{-1} \hat{X} (\hat{X} \hat{S}_\lambda^{-1} \hat{X}')^{-1} \hat{S}_\lambda^{-1}) \bar{u} \end{aligned} \quad (5.19)$$

where  $\phi$  and  $\hat{\xi}$  are vectors of elements of  $\mathcal{H}$  with  $\phi$  as in (4.9) and

$$\hat{\xi} = (\hat{\xi}_1, \hat{\xi}_2, \dots, \hat{\xi}_n)$$

with  $\hat{\xi}_i$  given by (5.3). By Lemma 2.7,  $\hat{z}_\lambda$  is the solution to the problem: Find  $z \in \mathcal{H}$  to minimize

$$\sum_{i=1}^n (u_i - \langle \hat{\eta}_i, z \rangle)^2 + \lambda \int_0^1 (L_m z)^2 ds \quad (5.20)$$

where

$$\langle \hat{\eta}_i, z \rangle = \sum_{k=1}^N K(t_i, s_k) \omega_k z(s_k), \quad i = 1, 2, \dots, n$$

and  $\lambda$  is of the order of magnitude of the mean square quadrature error. Thus, if the primary source of error in forming a computational estimate of  $z(s)$  is quadrature error, then this shows that an appropriate choice for the regularizing parameter  $\alpha$  of (1.6) is as  $\lambda$ , an estimate of the mean square quadrature error, as defined by (5.10) and (5.16).

We have not mentioned the choice of quadrature formula. Once the quadrature points  $\{s_k\}_{k=1}^N$  are chosen, the choice of best formula in the sense of Sard is equivalent to approximating the element  $\eta_i \in \mathcal{H}$  by a linear estimation of the elements  $\{R_{s_k}\}_{k=1}^N$ . The optimum coefficients in this case

are readily seen to depend on the unknown integrals. Hence a convenient quadrature formula which allows a "guestimate" of  $\lambda$  should be used.

As is widely known, as soon as there are experimental or computational errors, there is a point of diminishing returns in choosing  $n$  too large. If, e.g.  $K(t,s)$  is continuous then  $||\hat{\Sigma}^{-1}||$  and  $||\hat{\Sigma}^{-1}|| \rightarrow \infty$  as  $n$  becomes large, where  $||\cdot||$  is the spectral norm. We will quantify this statement and indicate a mitigating technique. To simplify the equations, we let  $M_v z = 0$ ,  $v = 1, 2, \dots, m$ . Then we may let

$$u_i = \langle \hat{\xi}_i + \epsilon_i, z \rangle = \int_0^1 K(t_i, s) z(s) ds, \quad i = 1, 2, \dots, n \quad (5.21)$$

where  $\{\hat{\xi}_i\}_{i=1}^n$  are given by (5.3) and  $\epsilon_i = \xi_i - \hat{\xi}_i$ . Then, by Theorem 4.2 the solution to the problem: Find  $z \in H$  to minimize

$$\sum_{i=1}^n (u_i - \langle \hat{\xi}_i, z \rangle)^2 + \lambda \int_0^1 (L_m z)^2 ds \quad (5.22)$$

is given by  $\hat{z}_\lambda$ ,

$$\hat{z}_\lambda = \hat{\xi} (\hat{\Sigma} + \lambda I)^{-1} \bar{u} \quad (5.23)$$

and

$$|z(s) - \hat{z}_\lambda(s)|^2 = |\langle z - \hat{z}_\lambda, R_{1,s} \rangle| =$$

$$|\langle z, R_{1,s} - \tilde{R}_{1,s,\lambda} \rangle| \leq \|z\| \|R_{1,s} - \tilde{R}_{1,s,\lambda}\| \quad (5.24)$$

where

$$\tilde{R}_{1,s,\lambda} = \sum_{i=1}^n d_i (\hat{\xi}_i + \epsilon_i)$$

with  $\bar{d} = (d_1, d_2, \dots, d_n)$  given by

$$\bar{d} = \hat{\xi}(s) (\hat{\Sigma} + \lambda I)^{-1}$$

with  $\hat{\xi}(s)$  given by (5.14b).

Now

$$\|R_{1,s} - \tilde{R}_{1,s,\lambda}\|^2 \leq 2\{\|R_{1,s} - \sum_{i=1}^n d_i \hat{\xi}_i\|^2 + \|d_i \epsilon_i\|^2\}$$

(5.25)

If  $V$  is the matrix with  $i, j^{\text{th}}$  entry  $v_{ij} = \langle \epsilon_i, \epsilon_j \rangle$ , then the term in brackets in (5.25) may be expanded as

$$\begin{aligned}
R_1(s, s) &= 2\bar{\xi}(s)(\hat{\Sigma} + \lambda I)^{-1}\bar{\xi}(s)' + \bar{\xi}(s)(\hat{\Sigma} + \lambda I)^{-1}\hat{\Sigma}(\hat{\Sigma} + \lambda I)^{-1}\bar{\xi}(s)' \\
&\quad + \bar{\xi}(s)(\hat{\Sigma} + \lambda I)^{-1}V(\hat{\Sigma} + \lambda I)^{-1}\bar{\xi}(s)' \quad (5.26)
\end{aligned}$$

and some algebraic manipulation gives that (5.26) is equal to

$$R_1(s, s) = \bar{\xi}(s)\hat{\Sigma}^{-1}\bar{\xi}(s)' + \bar{\xi}(s)\hat{\Sigma}^{-1/2}A\hat{\Sigma}^{-1/2}\bar{\xi}(s)' \quad (5.27)$$

where  $\hat{\Sigma}^{-1/2}$  is the symmetric square root of  $\hat{\Sigma}^{-1}$  and

$$A = (\hat{\Sigma} + \lambda I)^{-1}(\hat{\Sigma}^{1/2}V\hat{\Sigma}^{1/2} + \lambda^2 I)(\hat{\Sigma} + \lambda I)^{-1} \quad (5.28)$$

Letting  $P_{\hat{\xi}}$  denote the projection operator onto the subspace spanned by  $\{\hat{\xi}_i\}_{i=1}^n$ , we have

$$\|P_{\hat{\xi}}R_{1,s}\|^2 = \bar{\xi}(s)\hat{\Sigma}^{-1}\bar{\xi}(s)' \quad (5.29)$$

Lines (5.25), (5.27), (5.28) and (5.29) yield the bound

$$\begin{aligned}
\|R_{1,s} - \tilde{R}_{1,s,\lambda}\|^2 &\leq 2 \left[ \|R_{1,s} - P_{\hat{\xi}}R_{1,s}\|^2 + \right. \\
&\quad \left. + \|P_{\hat{\xi}}R_{1,s}\|^2 \cdot \|(\hat{\Sigma} + \lambda I)^{-1}\| \cdot \|V + \lambda I\| \right] \quad (5.30)
\end{aligned}$$

If  $V = \lambda I$ , the term in curly brackets in (5.30) can be replaced by  $\lambda \|(\hat{\Sigma} + \lambda I)^{-1}\|$ . In practice as  $N$ , the number of quadrature

points becomes large, one expects  $||V||$  to decrease to a finite limit imposed by round off errors. As  $n(n \ll N)$ , the number of data points, becomes large,  $||R_{1,s} - P_{\hat{\xi}} R_{1,s}||$  decreases, however,  $\lambda ||(\hat{\Sigma} + \lambda I)^{-1}||^2$  will increase for  $\lambda$  bounded below since  $||(\hat{\Sigma})^{-1}|| \rightarrow \infty$  as  $n \rightarrow \infty$ .

Let  $H = \{h_{li}\}$  be a  $p \times n$  matrix of real numbers of rank  $p$ . Let

$$\begin{aligned}\hat{\xi}_l &= \sum_{i=1}^n h_{li} \hat{\xi}_i \\ \hat{\epsilon}_l &= \sum_{i=1}^n h_{li} \epsilon_i \\ \underline{u}_l &= \sum_{i=1}^n h_{li} u_i = \sum_{i=1}^n h_{li} \int_0^1 K(t_i, s) z(s) ds, \quad l = 1, 2, \dots, p.\end{aligned}\tag{5.31}$$

If we estimate  $z(s)$  assuming the information  $z \in H_1$ ,

$$\langle z, \hat{\xi}_l + \hat{\epsilon}_l \rangle = \underline{u}_l, \quad l = 1, 2, \dots, p$$

as before, only replacing the set  $\{\hat{\xi}_i\}_{i=1}^n$  by  $\{\hat{\xi}_l\}_{l=1}^p$  and the numbers  $\{u_i\}_{i=1}^n$  by  $\{\underline{u}_l\}_{l=1}^p$ , an error bound analogous to the right hand side of (5.30) is obtained of the form

$$2(||R_{1,s} - P_{\hat{\xi}} R_{1,s}||^2 + ||P_{\hat{\xi}} R_{1,s}||^2 \{ ||(H \hat{\Sigma} H' + \lambda I)^{-1}|| \cdot ||H V H' + \lambda I|| \})\tag{5.32}$$



where  $P_{\hat{\xi}}$  is the projection operator onto the subspace spanned by  $\{\hat{\xi}_\ell\}_{\ell=1}^p$ . If the rows of  $H$  are orthonormal vectors in Euclidean  $n$ -space, then the term in curly brackets in (5.32) can only be decreased, as compared to the term in brackets in (5.20), while the first term in (5.32) may, for large  $n$ , and  $n-p$ , may not increase much, as compared to the first term in (5.30). If  $V = \lambda I$ , then the "optimum" choice of  $H$  to minimize the term in brackets in (5.32) is to choose the rows of  $H$  as the eigenvectors corresponding to the  $p$  largest eigenvalues of  $\hat{\xi}$ .

#### 6. 'Statistical Estimation' of Solutions to Integral Equations

It is far from coincidental that the method of regularization and the method of statistical estimation lead to the same numerical solution. Let  $R(s, s')$  be a continuous positive definite kernel on  $S \times S$ , and  $\mathcal{H}_R$  the reproducing kernel Hilbert space associated with  $R$ . Let  $Z(s)$ ,  $s \in S$  be a stochastic process (i.e. a family of random variables indexed by  $s$ ), with  $EZ(s) = 0$  2] and

$$EZ(s)Z(s') = R(s, s') \quad (6.1)$$

2] without real loss of generality

Let  $K(t,s)$  satisfy the assumptions of the corollary to Theorem 1.1 and consider the stochastic model

$$U(t) = \int_S K(t,s)Z(s)ds \quad (6.2)$$

where observations will be taken on the random variables  $U(t_i)$ ,  $i = 1, 2, \dots, n$ .  $U(t)$  is a well defined random variable for every  $t \in T$ . We have been studying the (deterministic) model

$$u(t) = \int_S K(t,s)z(s)ds \quad (6.3)$$

where  $z \in \mathcal{H}_R$ , and the numbers  $u(t_i)$ ,  $i = 1, 2, \dots, n$  are available. The purpose of this section is to demonstrate rigorously that the same numerical solution to the integral equation is obtained whether the true 'solution' is considered to be an element  $z \in \mathcal{H}_R$  or a realization of a stochastic process  $Z(s)$  with covariance  $R(s,s')$ .

Let  $\mathcal{H}_Z$  be the Hilbert space spanned by the stochastic process  $\{Z(s), s \in S\}$ . (See [6]).  $\mathcal{H}_Z$  is defined as follows: All random variables  $Y$  which are finite linear combinations of the form

$$Y = \sum_{\ell} a_{\ell} Z(s_{\ell}) \quad s_{\ell} \in S \quad (6.4)$$

are in  $\mathcal{H}_Z$ . An inner product on the linear manifold of all

such finite linear combinations is

$$\langle Y_1, Y_2 \rangle = E Y_1 Y_2, \quad (6.5)$$

and  $\mathcal{H}_Z$  is the closure of this linear manifold with the given inner product. The precise source of the duality between 'deterministic' and statistical models is the following (well known) fact:

$\mathcal{H}_Z$  is isometrically isomorphic to  $\mathcal{H}_R$  under the isomorphism induced by the correspondence " $\sim$ ",

$$Z(s) \sim R_s, \quad \forall s \in S \quad (6.6)$$

Furthermore, the random variable  $Y \in \mathcal{H}_Z$  corresponds to the element  $\eta \in \mathcal{H}_R$  if and only if

$$E Y Z(s) = \langle \eta, R_s \rangle = \eta(s), \quad s \in S \quad (6.7)$$

The family of random variables  $U(t)$ ,  $t \in T$  are all in  $\mathcal{H}_Z$  by our assumptions on  $K(t, s)$  and

$$U(t) \sim \eta_t \quad (6.8)$$

where  $\eta_t$  is defined, for each  $t$ , by (1.25). Let  $\{\phi_v\}_{v=1}^m$  be a specified set of  $m$  orthonormal elements in  $\mathcal{H}_R$ , and let

$\{\rho_v\}_{v=1}^m$  be the  $m$  random variables in  $\#_Z$  which correspond to  $\{\phi_v\}_{v=1}^m$  under the isomorphism induced by (6.6). It can then be shown that  $Z(s)$ ,  $s \in S$  has a representation of the form

$$Z(s) = \sum_{v=1}^m \rho_v \phi_v(s) + Z_1(s) \quad (6.9)$$

where  $Z_1(s)$  is a zero mean stochastic process with

$$EZ_1(s)Z_1(s') = R_1(s, s') = R(s, s') - \sum_{v=1}^m \phi_v(s)\phi_v(s') \quad (6.10)$$

and  $E\rho_v\rho_\mu = \delta_{\mu,v}$ ,  $\mu, v = 1, 2, \dots, m$ . Let  $\hat{Z}(s)$  be, for each  $s$ , that random variable in the subspace of  $\#_Z$  spanned by  $\{U(t_i)\}_{i=1}^n$  which minimizes

$$E(\hat{Z}(s) - Z(s))^2 \quad (6.11)$$

subject to

$$E(\hat{Z}(s) - Z(s) | \rho_v, v = 1, 2, \dots, n) = 0. \quad (6.12)$$

It follows from Lemma 2.2 and (6.6) with the identifications  $Z(s) \sim \delta_0$ ,  $U(t_i) \sim \eta_i$ ,  $\rho_v \sim \phi_v$ , and  $\int K(t_i, s)Z_1(s)ds \sim \xi_i$ , that

$$\begin{aligned}\hat{Z}(s) &= \phi(s) (X\bar{\Gamma}^{-1}X')^{-1}X\bar{\Gamma}^{-1}U \\ &+ \xi(s) (\bar{\Gamma}^{-1} - \bar{\Gamma}^{-1}X(X\bar{\Gamma}^{-1}X')^{-1}X\bar{\Gamma}^{-1})U\end{aligned}\quad (6.13)$$

where

$$\phi(s) = (\phi_1(s), \phi_2(s), \dots, \phi_m(s))$$

$$\xi(s) = (\xi_1(s), \xi_2(s), \dots, \xi_n(s))$$

$X$  and  $\bar{\Gamma}$  are as in (2.1), and  $U$  is the vector of random variables given by

$$U = (U(t_1), U(t_2), \dots, U(t_n))$$

Thus, the numerical value  $\hat{Z}(s)$  based on the model (6.2) and a "realization"  $U(t_i) = u_i$ ,  $i = 1, 2, \dots, n$ , is exactly the same as the numerical value of  $\hat{z}_0(s)$  of (4.11), based on the model (6.3). An identical statement may be made about  $\hat{z}_\lambda(s)$  of (5.19) if we replace (6.2) by the stochastic model

$$U(t_i) = \sum_{k=1}^N K(t_i, s_k) \omega_k Z(s_k) + \epsilon(t_i) \quad i = 1, 2, \dots, n \quad (6.14)$$

where

$$E(t_i) = 0$$

$$E(t_i)E(t_j) = 0, \quad i \neq j$$

$$= \lambda, \quad i = j$$

$$E(t_i)Z(s_k) = 0, \quad i = 1, 2, \dots, n$$

$$k = 1, 2, \dots, N$$

If we let  $H_R^*$  be the Hilbert space of all continuous linear functionals on  $H_R$ , then  $H_R^*$  is consequently isometrically isomorphic to  $H_Z$  under the correspondance

$$Z(s) \sim N_s$$

where  $N_s$  is the continuous linear functional defined by

$$N_s z = z(s), \quad z \in H_R$$

Then

$$U(t_i) \sim \Lambda_{t_i}$$

where  $\Lambda_{t_i}$  is the continuous linear functional defined by

$$\Lambda_{t_i} z = \int_S K(t_i, s) z(s) ds, \quad z \in H_R$$

It is seen that the geometry for approximating  $Z(s)$  by  $\{U(t_i)\}_{i=1}^n$  is exactly the same as the geometry for approximating  $N_s$  by  $\{\Lambda_{t_i}\}_{i=1}^n$ .

An experimenter approaching the problem with the model (6.2) chooses the prior covariance  $R(s, s')$  of (6.1) according to his belief or past experience concerning  $Z(s)$ . The numerical analyst, beginning with (6.3) should choose an  $R$  such that the norm of the solution  $z$  in  $H_R$  is known or believed to be small.

It is clear that algorithms for the numerical solution of a broad variety of (linear) equations can, in fact, be identified with prediction problems on stochastic processes in this manner.

## REFERENCES

1. Ahlberg, T.H., Nilson, E.W., and Walsh, J.L., The Theory of Splines and Their Applications, (1967), Academic Press, New York.
2. Aronszajn, N., Theory of reproducing kernels, Trans. Am. Math. Soc., 68 (1950), pp. 337-404.
3. de Boor, C. and Lynch, R.E., On splines and their minimum properties, J. Math. Mech., 15 (1967), pp. 953-69.
4. Golomb, M. and Weinberger, H.F., Optimal approximation and error bounds, pp. 117-90 of On Numerical Approximation (R.E. Langer, Ed.), University of Wisconsin Press, Madison, Wisconsin (1959).
5. Mangasarian, O.L., and Schumaker, L.L., Splines via optimal control, to appear in the proceedings of the Conference on Approximations, with Special Emphasis on Splines, held at the University of Wisconsin, May 4-6, 1969.
6. Parzen, E., An approach to time series analysis, Ann. Math. Statist., 32, (1961), pp. 951-989.
7. Phillips, David L., A technique for the numerical solution of certain integral equations of the first kind, J. Assn. Comput. Mach. 9 (1962), pp. 94-97.
8. Ribiere, G., Regularisation d'Operateurs, R.I.R.O. 1, 5, (1967), pp. 57-79.
9. Sard, A., Linear Approximation, (Mathematical Surveys, No. 9), American Mathematical Society, Providence, R. I. (1963).
10. Slepian, D. and Pollak, H.O., Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty - I, Bell System Technical Journal, 40, (1961), pp. 43-64.
11. Strand, O.N. and Westwater, E.R., Statistical estimation of the numerical solution of a Fredholm integral equation of the first kind, J.A.C.M. 15, 1, (1968), pp. 100-114.



12. Tihonov, A.N., Solution of incorrectly formulated problems and the regularization method, Soviet Mathematics 4, #4, (1963), pp. 1035-38.
13. Tihonov, A.N., Regularization of incorrectly posed problems, Soviet Mathematics 4, #6, (1963), pp. 1624-27.
14. Tihonov, A.N. and Glasko, V.B., An approximate solution of Fredholm integral equations of the first kind, Ž. Vyčisl. Mat. i Mat. Fiz. 4 (1964), pp. 564-571.

DOCUMENT CONTROL DATA - R & D		
<small>(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)</small>		
1. ORIGINATING ACTIVITY (Corporate author) Mathematics Research Center, U. S. Army University of Wisconsin, Madison, Wis. 53706		2a. REPORT SECURITY CLASSIFICATION Unclassified
		2b. GROUP None
3. REPORT TITLE ON THE APPROXIMATE SOLUTION OF FREDHOLM INTEGRAL EQUATIONS OF THE FIRST KIND		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Summary Report: no specific reporting period.		
5. AUTHOR(S) (First name, middle initial, last name) Grace Wahba		
6. REPORT DATE May 1969	7a. TOTAL NO. OF PAGES 64	7b. NO. OF REFS 14
8a. CONTRACT OR GRANT NO. Contract No. DA-31-124-ARO-D-462	9a. ORIGINATOR'S REPORT NUMBER(S) #990	
8b. PROJECT NO. None	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) None	
10. DISTRIBUTION STATEMENT Distribution of this document is unlimited.		
11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Army Research Office-Durham, N.C.	
13. ABSTRACT <p>We treat the problem of obtaining good approximate solutions to</p> $u(t) = \int_S K(t,s) z(s) ds, \quad t = t_1, t_2, \dots, t_n,$ <p>as a problem in the approximation of one continuous linear functional in a reproducing kernel Hilbert space, by several others. The methods of regularization and statistical estimation are shown to be special cases. Pointwise error bounds and a criteria for choosing the regularization parameter are given.</p>		

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

AGO 1698A